11-2019

The chilling effect of enforcement of computer misuse: Evidences from online hacker forums

Qiu-hong WANG
Singapore Management University, qiuhongwang@smu.edu.sg

Rui-Bin GENG

Seung Hyun KIM

School of
Information Systems

SMU
SINGAPORE MANAGEMENT UNIVERSITY

The Chilling Effect of Enforcement of Computer Misuse: Evidences from Online Hacker Forums

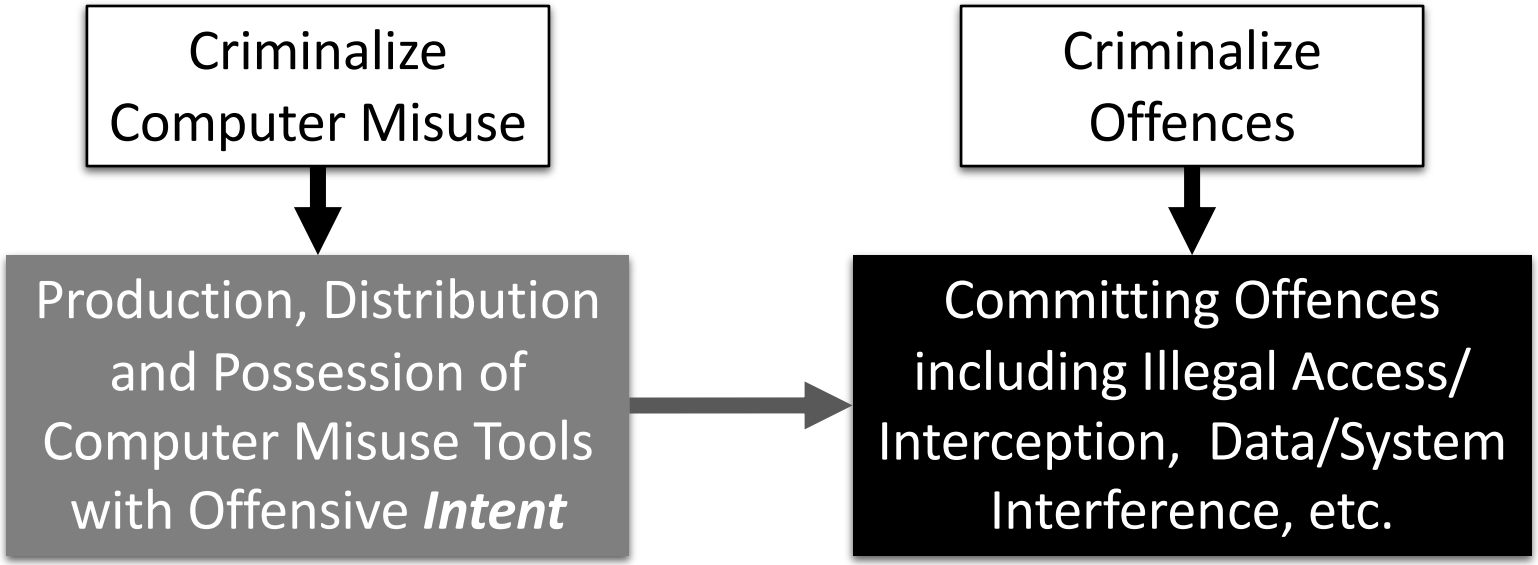
Assistant Professor: Qiu-Hong WANG

Singapore Management University

Co-authors: Rui-Bin Geng, Seung Hyun Kim

11 July 2019, Cambridge

Motivation -- Deterrence's Difficulty



Judicial determination of CM: the legal system need to take ex-ante adjudication of the malice and severity of hacking tools which may vary with the context where the act will be committed.

Perpetration cost advantage: Automation and reachability
Lower knowledge barrier to acquire hacking techniques via online communities

Enforcement cost disadvantage: Invisibility and anonymity
Jurisdictional boundary

List of Computer Misuse Act (CMA)

Country	Law	Amendment
Australia	Criminal Code Act 1995 (Cth)	ss 478.3 and 478.4
Croatia	New Criminal Law	Article 272
Canada	Protecting Canadians from Online Crime Act	Section 342.2
China	Criminal Code	Article 285
Colombia	Penal Code Act 1273 of 2009	Article 269A-J
Ethiopia	Telecom Fraud Offence Proclamation	Article 3
Fiji	Crimes Decree 2009	Article 346
France	Monetary and Financial Code	Article L163-4
Germany	German Criminal Code	Acts 202c
Italy	Penal Code	Art 615
Netherlands	Dutch Criminal Code	Article 350a
New Zealand	Crimes Amendment Act 2013 (2013 No 27)	subsection 1 of 251
Qatar	Cybercrime Law (No. 14 of 2014)	Article 66
Russia	Criminal Code	Act 273 and 138.1
Serbia	Criminal Code	Article 304a
Singapore	Computer Misuse and Cybersecurity Act	Article 10(1)
Sweden	Criminal Code	Article 9b
Switzerland	Criminal Code	Article 143bis
United Kingdom	Computer Misuse Act (UK)	s1, s3, s3A and s3ZA
United States	Computer Fraud and Abuse Act of 1986	(a)(5)(A)

UK: The Computer Misuse Act 1990: Section 3A: Making, supplying or obtaining articles for use in an offence under Section 1,3 or 3ZA

A person is guilty of an offence if he:

A. Makes, adapts, supplies or offers to supply any article intending it to be used to commit or to assisting the commission of an offence contained elsewhere in the Act
B. Supplies or offers to supply any article believing that it is likely to be used to commit or to assisting the commission of an offence contained elsewhere in the Act

US: The Computer Fraud and Abuse Act: (a)(5)(A)

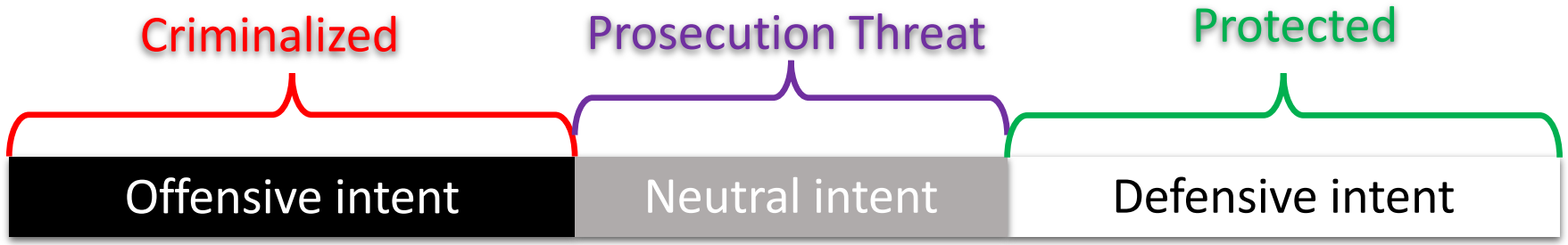
Knowingly causes the transmission of a program, information, code, or command, and as a result of such conduct, intentionally causes damage without authorization, to a protected computer

China: Criminal Code: the Amended Article 285

Any person who provides programs or instruments used specially for invading or illegally controlling computer information systems, or knowingly provides programs or instruments to another person for committing illegal or criminal acts of invading or illegally controlling computer information systems shall, if the circumstances are serious, be punished in accordance with the provisions of the preceding paragraph

Motivation – Deterrence or Chilling Effect?

- My python password finder for any site!
- Easily Hackable important Website :)
- [The Order] Free Rat Support | Reliable | Quick and Easy | 2+ Years of Experience
- Hacking A College
- DDoS Service [Cheap] [Powerful]
- Ten to fifteen thousand proxies in a list [ip:port].
- How to change your ip in less then 1 minute
- Anonymity complete GUIDE By Theraider & Dangerous R.
- Ping Scan Script
- How to Know when you are infected with RATs or Keyl0ggers.
- How to protect your HTML source code
- How to stop people from resolving your IP via Skype
- Nexus anti-flood 2010 with DDOS protection!



Judicial determination of CM:

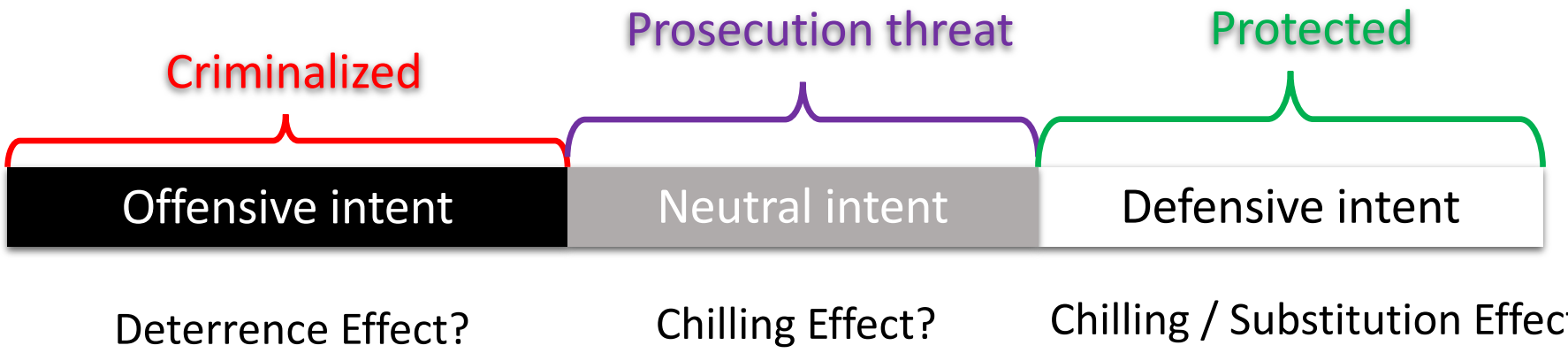
- Legal system with fallibility and uncertainty
- Predict potential cybersecurity risks associated with new technology or new uses of existing technology
- Dual use nature of cybersecurity technology: tools for penetration tests; cryptocurrency
- Unfalsifiability of security claims

Motivation – Concerns on Chilling Effect

- Cost of Chilling Effect:
  - Defamation vs. Free Speech
  - Government surveillance vs. Privacy
  - Cybersecurity Offense vs. Defense
- Empirical Challenge of Chilling effects
  - Where to find a control group?
  - Lack of individual-level data to track a choice between different intents
  - Globalized activities
  - Shift in norms

Research Questions -- Empirical Evidence of Chilling Effect

- **External Shock**: CMA enforcement -- the production, distribution, and possession of hacking tools with offensive *intent*
- **Context**: Publicly accessible online hacker communities



- While the CMA enforcement explicitly imposes legal risk on the communication with *offensive intent*, would the supposition of this deterrence effect lead to the chilling effect on the sharing with *neutral intent or even defensive intent*?
- How would the online social community context *reinforce* or *weaken* the effects of CMA enforcement?

Research Context -- Hacker Forums operated in the surface web as vantage points for diversified intents

- Moral ambiguity leads to the coexistence of black/grey/white hats in online hacker communities, and discussions on offense, defense or neutral-intent techniques with dual use (Thomas 2005)
- Dual roles
  - A stepping stone towards more serious online cyber-attacks (Pastrana et al. 2018)
  - A school for white hats and grey hats to understand hacking techniques (Kirsch 2014).
- Not for the most malicious activities but less determined hackers or the curious (Pastrana et al. 2018)

Research Context -- Chinese Hacker Forums

- CMA enforcement -- February 28, 2009, the Amendment of Article 285 in the Criminal Law
- Language barrier and Internet access filtering lead to localized subjects and their limited mobility
  - hackforums.net was not accessible in China
  - The earliest Chinese dark web was launched in October 2014
- Two top forums
  - Ranked the 2nd and 3rd (Alexa.com →China→ Computers/Security → Hacker , April 05, 2017)
  - 89.4%~92.6% of the users geographically located in China

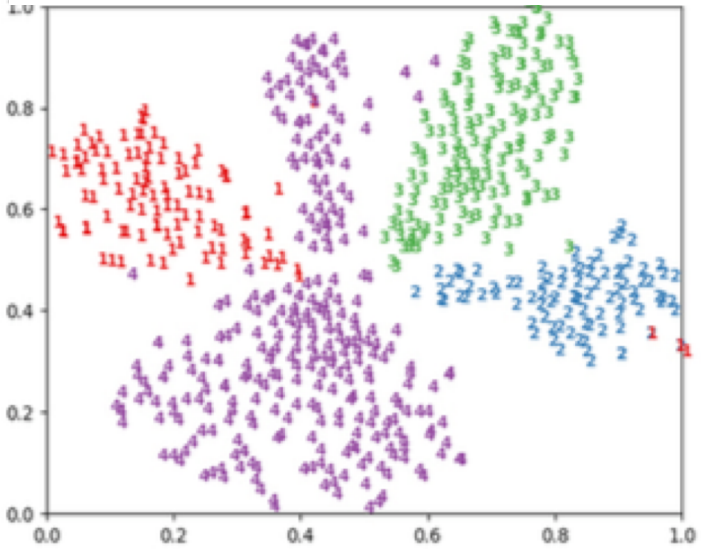
-- The majority of the forum participants are within the jurisdictional scope of the CMA enforcement

Context and Data: Author Intent Classification

Exploratory knowledge
4 word embedding clusters →
4 categories of contribution intents

The training and testing datasets

Unsupervised Clustering based on semantic cliques



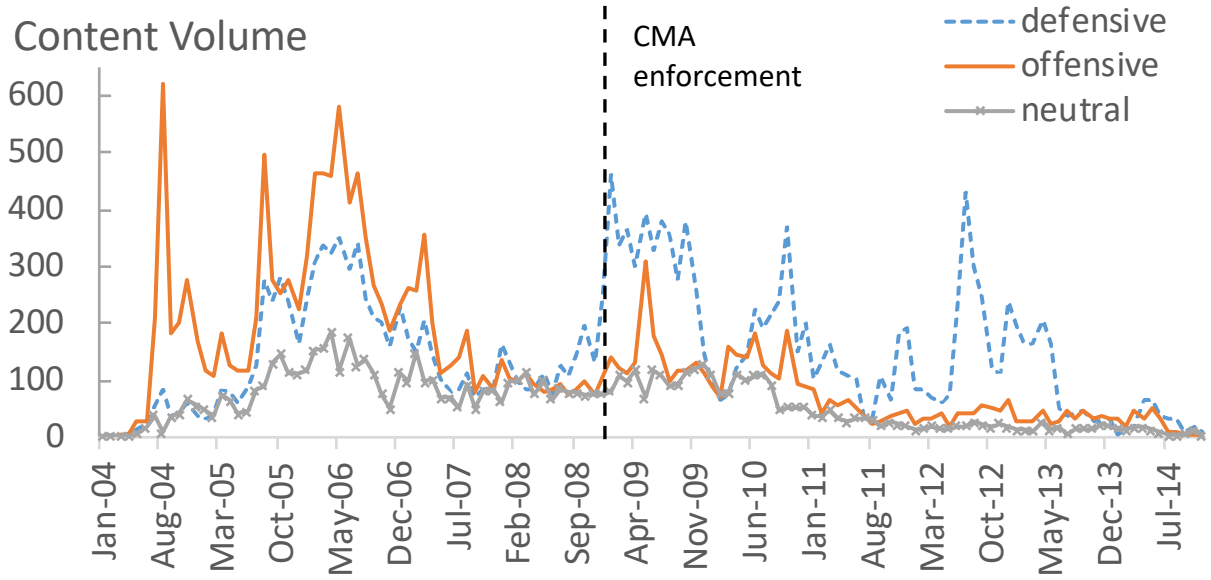
Manual Labelling

- Two human coders after 6 months of training
- 25% of leading posts in each year:
  - Forum A: 38,736 / 165,870
  - Forum B: 12,093 / 52,154
- 50,827 consistently labelled records
- inter-rater agreement: 0.87 for Forum A and 0.92 for Forum B

NLP-CNN model

	precision	recall	F1
irrelevant	0.98	0.99	0.98
defensive	0.95	0.94	0.94
offensive	0.96	0.93	0.95
neutral	0.94	0.90	0.92

Preliminary Analysis



	Before Enforcement	After Enforcement
The number of leading posts	137,718	80,306
The number of replies per leading post	7.53	10.09
% of defensive leading posts	6.62%	12.63%
% of offensive leading posts	8.78%	5.84%
% of neutral leading posts	3.67%	3.59%
% of irrelevant leading posts	80.97%	77.86%

Quasi-Difference-In-Difference

A reduced-form regression on the number of posts in different categories {defensive, offensive, neutral, irrelevant} generated by hacker forum user i in month t (Marthews & Tucker 2017)

$$\ln(Postcnt_{ijt}) = \beta_{0i} + \beta_{1i}\ln(Postcnt_{ij,t-1}) + \beta_{2i}AfterCMA_t + \beta_{3i}Category_{it}$$
$$+\beta_{4i}AfterCMA_t \times Category_{it} + \beta_{5i}\ln(TotalPost_{i,t-1})$$
$$+\beta_{6i}Age_{it} + \beta_{7i}Age_{it}^2 + \beta_{8i}Month_t + \beta_{9i}Month_t^2 + \tau_t + u_i + e_{it}$$

$AfterCMA_t \times Offensive_{it}$	-0.0248*** (0.0002)
$AfterCMA_t \times Defensive_{it}$	0.0262*** (0.0002)
$AfterCMA_t \times Neutral_{it}$	-0.0273*** (0.0002)
Adjusted R-squared	0.1038
No. of observations	2,826,232

Limitations

- ❑ Inflation with many zero observations
- ❑ User's contribution intent decision interdependent within each user
- ❑ Contribution on security-irrelevant posts is correlated with security-related posts
- ❑ No way to address forum self-regulation on obviously illegal posts

A Mixed Nested Logit Model

Each choice occasion: whether to post and which to post

$$\tilde{U}^{A}_{ijk} = \beta_{0ij} + \beta_{1i} Age_k + \beta_{2i} Age_k^2 + \delta_{1j} AfterCMA_k + \delta_{2j} Experience_{ijk-1} + \delta_{3j} Attention_{ijk-1}$$
$$+ \delta_{4j} Peer_{ijk-1} + \delta_{5j} AfterCMA_k \bullet Experience_{ijk-1} + \delta_{6j} AfterCMA_k \bullet Attention_{ijk-1}$$
$$+ \delta_{7j} AfterCMA_k \bullet Peer_{ijk-1} + \delta_{8j} Other\_Post_{ijk-1}$$
$$= \beta_i X_{ik} + \delta_j W_{ijk}$$

Randomized heterogeneity across contributors on preference and life cycle

$$L(A_{ijk}, I_{it}) = \left(\frac{1}{1+\exp(\tilde{U}^{I}_{it})} \right)^{1-I_{it}} \prod_{j=1}^{J-1} \left[e^{-I_{j=2}\lambda\tau} \cdot \frac{\exp(\tilde{U}_{ijk})}{1+\sum_{j=1}^{J-1}\exp(\tilde{U}_{ijk})} \cdot \frac{\exp(\tilde{U}^{I}_{it})}{1+\exp(\tilde{U}^{I}_{it})} \right]^{A_{ijk} I_{it}}$$

Probability of being removed by forum self-regulation

Probability of {Offensive, Neutral, Defensive} post

Probability of post

	Delta (δ_j)	Main Effects	Interaction Effects
Offensive	AfterCMA	-0.2530***	-0.4292**
	Experience	0.0782***	0.0522***
	AfterCMA ×Experience		0.1153***
	Attention	0.0008**	0.0025**
	AfterCMA ×Attention		-0.0020***
	Peer	0.0898***	0.0889***
	AfterCMA×Peer		0.0111***
	ln(Other_Posts)	-0.6440***	-0.6434***
Defensive	AfterCMA	0.3308***	0.1953**
	Experience	0.0089***	0.0412***
	AfterCMA ×Experience		-0.0334***
	Attention	0.0322***	0.0338***
	AfterCMA ×Attention		0.0291***
	Peer	0.1305***	0.1117***
	AfterCMA×Peer		0.0955***
	ln(Other_Posts)	-0.5860***	-0.5126***
Neutral	AfterCMA	-0.1352***	-0.0801***
	Experience	0.0076	0.0870**
	AfterCMA ×Experience		-0.1832***
	Attention	0.0023**	0.0062***
	AfterCMA ×Attention		-0.0054***
	Peer	0.2349***	0.2421***
	AfterCMA×Peer		-0.0531**
	ln(Other_Posts)	-0.9735***	-0.9846***

	Avg. Marginal Effect on Probability	% Change in Probability
Deterrence Effect	-0.019***	↓21.87%
Substitution Effect	0.020***	↑28.71%
Chilling Effect	-0.004**	↓11.13%

Weakened

Reinforced

Weakened

- Diminishing marginal perpetration cost
- Increasing severity
- Increasing enforcement cost

Weakened

Reinforced

Reinforced

- Diminishing marginal utility
- Increasing utility
- Increasing utility

Reinforced

Reinforced

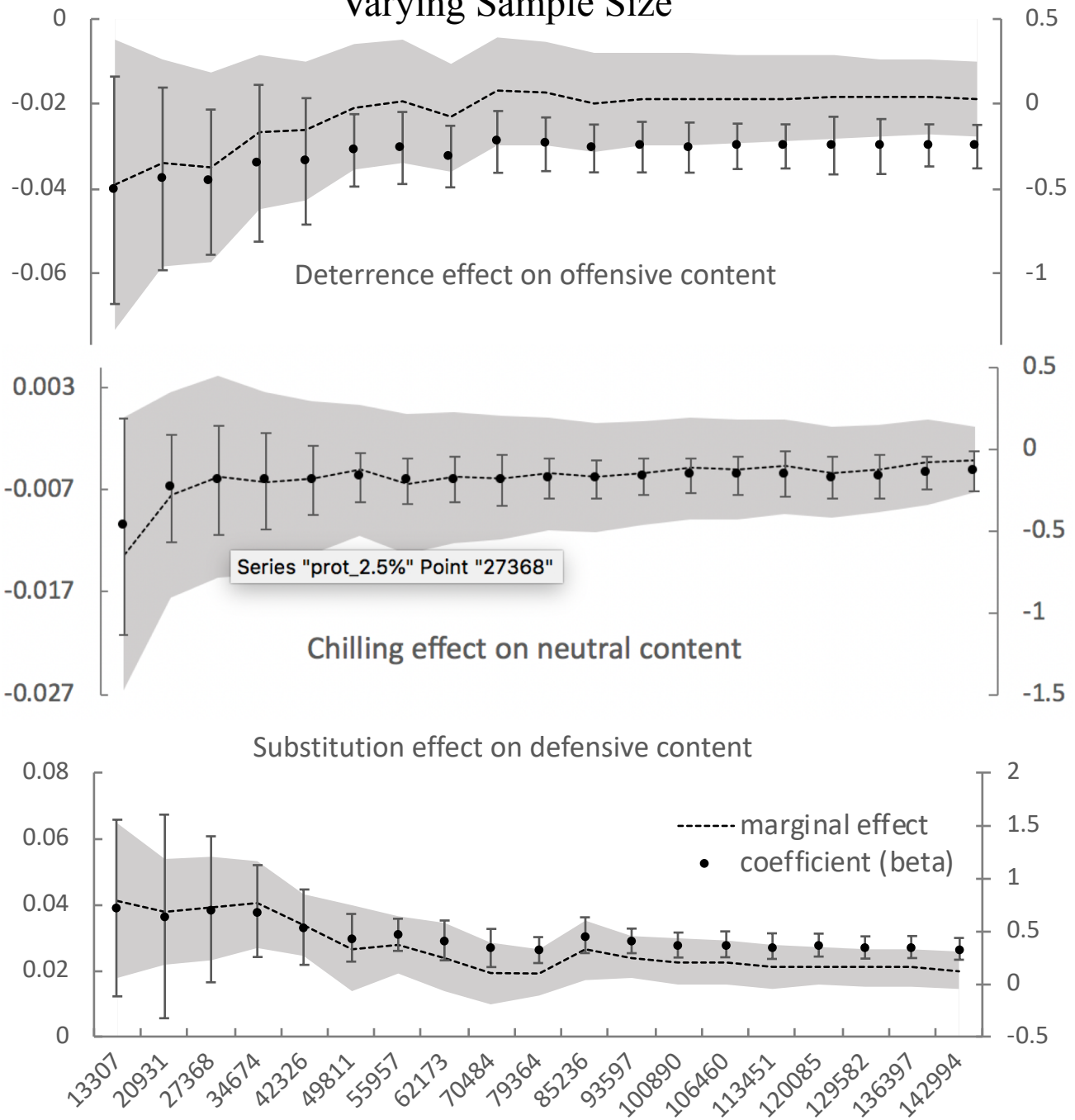
Reinforced

- Increasing probability of erroneous prosecution
- Exemplified perceived risk associated with social interaction (Kasperson et al. 1988)

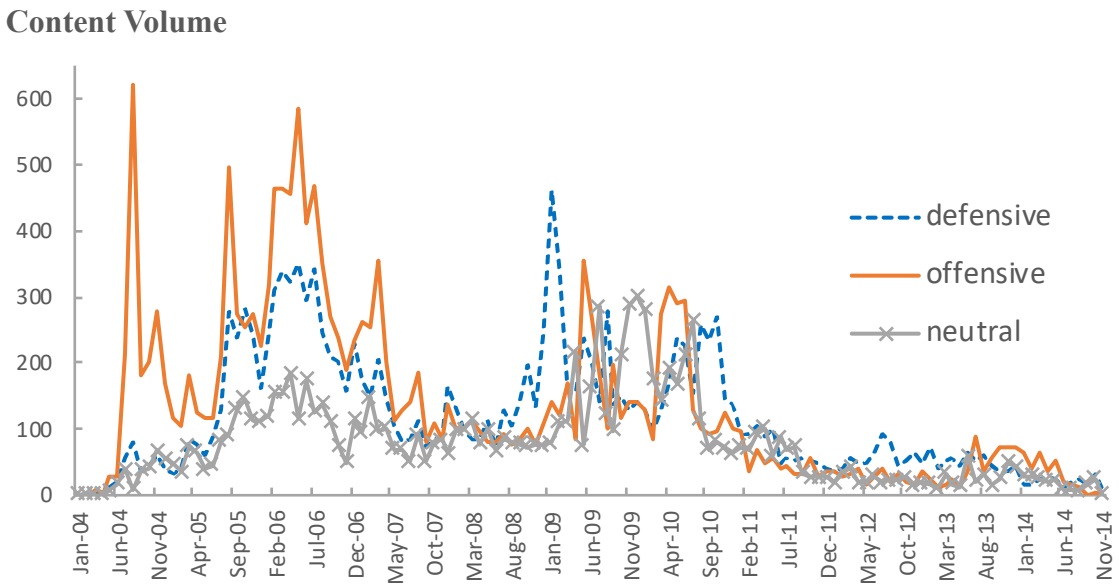
Robustness and Falsification Tests

- Subsamples by varying size or varying user activeness
- Alternative Models Fitness
- Alternative explanations related to
  - Competing peer forums (impacts on different contribution intent)
    - 3 major vulnerability disclosure forums
  - Shifting norms on forum users' topic preferences
    - Global or National Google Trends Index of 30 cybersecurity keywords
- If the enforcement is assumed six months in advance?
- If the enforcement did not occur at all?

Varying Sample Size



A Counterfactual Scenario without CMA Enforcement



Research Implications

- Initial empirical evidence of chilling effect of the CMA enforcement
- Chilling effect could be strengthened in online communities
- Domestic legislation may deter publicly-observable cybercrimes when the illegal activities are localized due to language barrier and internet accessibility control (Png et al. 2008)
- Deterrence effect may be weakened due to the diminishing marginal cost associated with experienced perpetrators and the increasing enforcement cost associated with the number of perpetrators (Katyal 1997)
- Positive substitution effect of domestic enforcement on promoting security defense as a result of the dual use nature of hacking techniques and the contribution incentives on the online social communities (Png et al. 2008)

Practical Implications

- The balance between deterrence and chilling effects
  - Cost disadvantage of the traditional security measures, e.g., investment and enforcement in tackling the never-ending cybersecurity risks
  - The importance of information sharing among the communities consisting of white/grey/black hats
- Forum administrators: feasible measures to promote the positive loop for cybersecurity in online hacker forums.
  - Increase public attention to both offensive posts and defensive posts. (Yue et al. 2019)
  - Increase the incremental benefit of defensive content contribution

SMU
SINGAPORE MANAGEMENT UNIVERSITY

Contact:

qiuhongwang@smu.edu.sg

Thank You!